Any Hermitian Matrix is a Linear Combination of Four Projections

Yoshihiro Nakamura
Division of Applied Mathematics
Research Institute of Applied Electricity
Hokkaido University, Sapporo, 060, Japan

Submitted by T. Ando

ABSTRACT

Any Hermitian matrix A is a linear combination of four projections. The number of projections can be reduced to three or two according as $\operatorname{rank}(A) \leqslant 7$ or $\leqslant 3$.

We work with the algebra $M_n(\mathbb{C})$ of complex $n \times n$ matrices. We will call a Hermitian idempotent a *projection* and think of $\mathbb{C}^n$ as the space of $n \times 1$ matrices. For $x \in \mathbb{C}^n$, we denote the conjugate transpose of x by x^*, which is a $1 \times n$ matrix. Every projection P of rank one in $M_n(\mathbb{C})$ has a form $P = xx^*$ with some unit vector $x \in \mathbb{C}^n$.

For each $n \in \mathbb{N}$, let $K(n)$ denote the minimum of all $k \in \mathbb{N}$ such that every $n \times n$ Hermitian matrix can be represented as a linear combination of k projections in $M_n(\mathbb{C})$. It is known that $K(n)$ is a number much smaller than n (see [1], [3], and [4]). Recently it was pointed out in [2] that $K(n) \leqslant 5$. In the present paper we shall prove $K(n) \leqslant 4$.

For $t \in \mathbb{R}$, let $[t]$ denote the integral part of t. Then our result is stated as follows.

THEOREM. $\min(3, [\log_2 n]+1) \leqslant K(n) \leqslant \min(4, [\log_2 n]+1)$.

As particular cases, we have the following.

COROLLARY. $K(1)=1$, $K(2)=K(3)=2$, *and* $K(4)=K(5)=K(6)=K(7) = 3$.

The proof of the theorem will be divided into several steps. Basic to the proof is the following elementary lemma on 2×2 matrices.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

LEMMA 1. *Given $\alpha, \beta, \lambda, \mu \in \mathbb{R}$, there exist projections P, Q of rank one in $M_2(\mathbb{C})$ such that*

$$\operatorname{diag}(\lambda, \mu) = \alpha P + \beta Q \tag{1}$$

if and only if

$$\lambda + \mu = \alpha + \beta \quad \text{and} \quad ||\alpha| - |\beta|| \leqslant |\lambda - \mu| \leqslant |\alpha| + |\beta|. \tag{2}$$

Proof. Since the general form of a projection of rank one in $M_2(\mathbb{C})$ is

$$\begin{bmatrix} \dfrac{1+s}{2} & \dfrac{\sqrt{1-s^2}}{2}e^{-i\theta} \\ \dfrac{\sqrt{1-s^2}}{2}e^{i\theta} & \dfrac{1-s}{2} \end{bmatrix} \quad \text{with} \quad s \in [-1,1] \text{ and } \theta \in \mathbb{R},$$

(1) holds for some P, Q if and only if there exist $s, t \in [-1, 1]$ and $\theta, \phi \in \mathbb{R}$ such that

$$\begin{aligned} \lambda &= \frac{\alpha(1+s)+\beta(1+t)}{2}, \\ \mu &= \frac{\alpha(1-s)+\beta(1-t)}{2}, \\ e^{i\theta}\alpha\sqrt{1-s^2} &+ e^{i\phi}\beta\sqrt{1-t^2} = 0. \end{aligned} \tag{3}$$

The simultaneous equations (3) are solvable if and only if

$$\begin{gathered} \lambda + \mu = \alpha + \beta, \qquad \lambda - \mu = \alpha s + \beta t, \\ \alpha^2(1-s^2) = \beta^2(1-t^2) \end{gathered} \tag{4}$$

hold for some $s, t \in [-1, 1]$. It is easily seen that the range of the continuous function $(s, t) \mapsto \alpha s + \beta t$, considered on the compact set $\{(s, t) \colon \alpha^2(1-s^2) = \beta^2(1-t^2),\ s, t \in [-1, 1]\}$, is the union of two intervals

$$\big[-|\alpha| - |\beta|, -||\alpha| - |\beta||\big] \cup \big[||\alpha| - |\beta||, |\alpha| + |\beta|\big].$$

Thus (2) is equivalent to (4). ∎

The lemma will be used in the following form.

LEMMA 2. *Let e_1, e_2 be unit vectors in $\mathbb{C}^n$ such that $e_1 = e_2$ or e_1 is orthogonal to e_2. If*

$$\alpha \geqslant \lambda \geqslant 0 \geqslant \mu \quad or \quad \lambda \geqslant \alpha \geqslant \mu \geqslant 0,$$

then there exist unit vectors $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{span}(e_1, e_2)$, the linear span of e_1, e_2, such that

$$\lambda e_1 e_1^* + \mu e_2 e_2^* = \alpha \boldsymbol{x}\boldsymbol{x}^* + (\lambda + \mu - \alpha)\boldsymbol{y}\boldsymbol{y}^*. \tag{5}$$

In fact, if $e_1 = e_2$, take $\boldsymbol{x} = \boldsymbol{y} = e_1$. Suppose that e_1 is orthogonal to e_2. If $\alpha \geqslant \lambda \geqslant 0 \geqslant \mu$, then (2) with $\beta = \lambda + \mu - \alpha$ is seen to be valid from the relations

$$|\alpha| + |\beta| = 2\alpha - \lambda - \mu \quad \text{and} \quad |\alpha| - |\beta| = \lambda + \mu,$$

while if $\lambda \geqslant \alpha \geqslant \mu \geqslant 0$, this follows from the relations

$$|\alpha| + |\beta| = \lambda + \mu \quad \text{and} \quad |\alpha| - |\beta| = 2\alpha - \lambda - \mu.$$

Thus Lemma 1 can be applied.

Now we proceed to the proof of the theorem.

Proof of $K(n) \leqslant [\log_2 n] + 1$. This follows from the following assertion:

Any Hermitian matrix A in $M_\ell(\mathbb{C})$ can be written as a linear combination of no more than $[\log_2\{\operatorname{rank}(A)\}] + 1$ projections in $M_\ell(\mathbb{C})$.

Let us prove the assertion by induction. The case $\operatorname{rank}(A) = 1$ is obvious because any matrix of rank one is a scalar multiple of a projection. Suppose that the assertion is true for all Hermitian matrices of rank not greater than $n - 1$. Let A be any Hermitian matrix of rank n, and represent it in the form

$$A = \sum_{j=1}^{n} \lambda_j e_j e_j^*, \tag{6}$$

where $\{e_j\}_{j=1}^n$ is an orthonormal system of eigenvectors, and the eigenvalues $\{\lambda_j\}_{j=1}^n$ are arranged in decreasing order

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n. \tag{7}$$

By considering $-A$ if necessary, it may be assumed that there is $k \in \mathbb{N}$ such that

$$k \geqslant n-k \quad \text{and} \quad \lambda_k > 0 > \lambda_{k+1} \tag{8}$$

(with the convention $\lambda_{n+1} = -\infty$).

First let us treat the case n is odd. Let $m = (2k-n+1)/2$. Then by (6)

$$A = \sum_{i=1}^{m-1} \{\lambda_{m-i} e_{m-i} e_{m-i}^* + \lambda_{m+i} e_{m+i} e_{m+i}^*\} + \lambda_m e_m e_m^*$$

$$+ \sum_{j=1}^{n-k} \{\lambda_{k+1-j} e_{k+1-j} e_{k+1-j}^* + \lambda_{k+j} e_{k+j} e_{k+j}^*\}.$$

When $k = n$ the last term does not appear. Let $\alpha = \lambda_m$. Since by (7) and (8)

$$\lambda_{m-i} \geqslant \alpha \geqslant \lambda_{m+i} \geqslant 0 \qquad (i = 1, 2, \ldots, m-1)$$

and

$$\alpha \geqslant \lambda_{k+1-j} \geqslant 0 \geqslant \lambda_{k+j} \qquad (j = 1, 2, \ldots, n-k),$$

according to Lemma 2 there exist unit vectors $\boldsymbol{x}_i, \boldsymbol{y}_i \in \operatorname{span}(e_{m-i}, e_{m+i})$ and $\boldsymbol{u}_j, \boldsymbol{v}_j \in \operatorname{span}(e_{k+1-j}, e_{k+j})$ such that

$$\lambda_{m-i} e_{m-i} e_{m-i}^* + \lambda_{m+i} e_{m+i} e_{m+i}^* = \alpha \boldsymbol{x}_i \boldsymbol{x}_i^* + (\lambda_{m-i} + \lambda_{m+i} - \alpha) \boldsymbol{y}_i \boldsymbol{y}_i^*$$
$$(i = 1, 2, \ldots, m-1)$$

and

$$\lambda_{k+1-j} e_{k+1-j} e_{k+1-j}^* + \lambda_{k+j} e_{k+j} e_{k+j}^* = \alpha \boldsymbol{u}_j \boldsymbol{u}_j^* + (\lambda_{k+1-j} + \lambda_{k+j} - \alpha) \boldsymbol{v}_j \boldsymbol{v}_j^*$$
$$(j = 1, 2, \ldots, n-k).$$

Since $\{\boldsymbol{x}_i, \boldsymbol{u}_j, e_m : i = 1, 2, \ldots, m-1, j = 1, 2, \ldots, n-k\}$ is an orthogonal family,

$$P = \sum_{i=1}^{m-1} \boldsymbol{x}_i \boldsymbol{x}_i^* + \sum_{j=1}^{n-k} \boldsymbol{u}_j \boldsymbol{u}_j^* + e_m e_m^*$$

is a projection, and $A = \alpha P + B$, where

$$B = \sum_{i=1}^{m-1} (\lambda_{m-i} + \lambda_{m+i} - \alpha) y_i y_i^* + \sum_{j=1}^{n-k} (\lambda_{k+1-j} + \lambda_{k+j} - \alpha) v_j v_j^*.$$

Since B is a Hermitian matrix and

$$\operatorname{rank}(B) \leqslant (m-1) + (n-k) = \frac{n-1}{2} = \left[\frac{n}{2}\right],$$

by the induction assumption it is a linear combination of no more than $[\log_2([n/2])] + 1$ projections. Now the induction is completed by the inequality

$$\left[\log_2\left(\left[\frac{n}{2}\right]\right)\right] + 1 \leqslant [\log_2 n].$$

Next, in the case where n is even, let $m = (2k-n)/2$. Then by (6)

$$A = \sum_{i=1}^{m} \left\{ \lambda_{m+1-i} e_{m+1-i} e_{m+1-i}^* + \lambda_{m+i} e_{m+i} e_{m+i}^* \right\}$$
$$+ \sum_{j=1}^{n-k} \left\{ \lambda_{k+1-j} e_{k+1-j} e_{k+1-j}^* + \lambda_{k+j} e_{k+j} e_{k+j}^* \right\}.$$

When $k = n$ the last term does not appear. Let $\alpha = \lambda_m$. If $2k = n$, the first term in the above does not appear and we take $\alpha = \lambda_1$. Just as in the above, there exist a projection P and a Hermitian matrix B such that

$$A = \alpha P + B \quad \text{and} \quad \operatorname{rank}(B) \leqslant m + (n-k) = \frac{n}{2};$$

hence the induction is completed analogously. ■

Proof of $K(n) \leqslant 4$. Let A be any $n \times n$ Hermitian matrix, and represent it in the form (6) with (7). By considering $-A$ if necessary, it may be assumed that

$$\sum_{j=1}^{n} \lambda_j \geqslant 0. \tag{9}$$

Set $\gamma_k = \Sigma_{j=1}^{k}\lambda_j$ for $k = 1,2,\dots,n$, $\varepsilon = 1/n\gamma_n$, and $\alpha = \max\{\gamma_k - (k-1)\varepsilon : k = 1,2,\dots,n\}$. Since by (7) and (9)

$$\gamma_k - (k-1)\varepsilon \geqslant 0 \geqslant -(\gamma_k - k\varepsilon) \qquad (k = 1,2,\dots,n),$$

according to Lemma 2 there exist unit vectors $\boldsymbol{x}_k, \boldsymbol{y}_k \in \operatorname{span}(\boldsymbol{e}_k, \boldsymbol{e}_{k+1})$ ($k = 1,2,\dots,n$) (with the convention $\boldsymbol{e}_{n+1} = \boldsymbol{e}_n$) such that

$$\{\gamma_k - (k-1)\varepsilon\}\boldsymbol{e}_k\boldsymbol{e}_k^* - (\gamma_k - k\varepsilon)\boldsymbol{e}_{k+1}\boldsymbol{e}_{k+1}^* = \alpha\boldsymbol{x}_k\boldsymbol{x}_k^* - (\alpha - \varepsilon)\boldsymbol{y}_k\boldsymbol{y}_k^*.$$

If $i \neq j$, then $\boldsymbol{x}_{2i}$ is orthogonal to $\boldsymbol{x}_{2j}$, and $\boldsymbol{x}_{2i-1}$ is orthogonal to $\boldsymbol{x}_{2j-1}$ (and similar relations hold for the family $\{\boldsymbol{y}_k\}_{k=1}^{n}$). Therefore all the matrices

$$P_1 = \sum_{k \text{ odd}} \boldsymbol{x}_k\boldsymbol{x}_k^*, \qquad P_2 = \sum_{k \text{ even}} \boldsymbol{x}_k\boldsymbol{x}_k^*,$$

$$Q_1 = \sum_{k \text{ odd}} \boldsymbol{y}_k\boldsymbol{y}_k^*, \qquad Q_2 = \sum_{k \text{ even}} \boldsymbol{y}_k\boldsymbol{y}_k^*$$

are projections, and

$$A = \sum_{k=1}^{n} \left(\{\gamma_k - (k-1)\varepsilon\}\boldsymbol{e}_k\boldsymbol{e}_k^* - (\gamma_k - k\varepsilon)\boldsymbol{e}_{k+1}\boldsymbol{e}_{k+1}^*\right)$$

$$= \alpha P_1 + \alpha P_2 - (\alpha - \varepsilon)Q_1 - (\alpha - \varepsilon)Q_2.$$

This proves $K(n) \leqslant 4$. ■

REMARK. In the above proof, the idea of gathering odd or even indexed vectors is due to [3], in which $K(n) \leqslant 6$ is proved.

Proof of $K(n) \geqslant 3$ for $n \geqslant 4$. Consider a matrix

$$A = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

such that

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0 \quad \text{and} \quad 2\lambda_1 > \sum_{j=1}^{n} \lambda_j. \tag{10}$$

Suppose that A is represented as a linear combination of two projections P, Q:

$$A = \alpha P + \beta Q \qquad (\text{with} \quad \alpha \geqslant \beta). \tag{11}$$

First let us show that $\beta > 0$. If $\beta \leqslant 0$, then the relation $\alpha P = A + (-\beta)Q$ implies that the projection P has rank n; hence $P = I$, the identity matrix. Then (11) implies that the number of distinct eigenvalues of A is at most 3, contradicting $n \geqslant 4$. Next, comparison of the traces of both sides of (11) yields

$$\sum_{j=1}^{n} \lambda_j = \alpha \operatorname{rank}(P) + \beta \operatorname{rank}(Q)$$

while $\lambda_1 \leqslant \alpha + \beta$ is an obvious consequence of (11) and $\alpha, \beta > 0$. When combined with (10), these inequalities imply

$$\alpha \operatorname{rank}(P) + \beta \operatorname{rank}(Q) < 2(\alpha + \beta),$$

which is possible only when $\operatorname{rank}(P) = 1$ and $\operatorname{rank}(Q) = n - 1$. Finally the relation

$$A - \beta I = \alpha P - \beta(I - Q)$$

leads to a contradiction, because

$$\operatorname{rank}(A - \beta I) \geqslant n - 1 \geqslant 3 \quad \text{but} \quad \operatorname{rank}(\alpha P - \beta(I - Q)) \leqslant 2.$$

This contradiction shows the impossibility of the representation (11); thus $K(n) \geqslant 3$. ■

The author wishes to thank Professor T. Ando for many helpful discussions.

REFERENCES

1 P. A. Fillmore, Sums of operators with square zero, *Acta Sci. Math.* (*Szeged*) 28:285–288 (1967).

2 K. Matsumoto, Self-adjoint operators as a real span of 5 projections, to appear.

3 A. Paszkiewicz, Any self-adjoint operator is a finite linear combination of projectors, *Bull. Acad. Polon. Sci. Sér. Sci. Math.* 28:337–345 (1980).

4 C. Pearcy and D. Topping, Sums of small numbers of idempotents, *Michigan Math. J.* 14:453–465 (1967).

Received 1 May 1983; revised 26 May 1983